

Figure/ground segregation from temporal delay is best at high spatial frequencies

著者	Kojima Haruyuki
journal or publication title	Vision Research
volume	38
page range	3729-3734
year	1998-01-01
URL	http://hdl.handle.net/2297/1250

doi: 10.1016/S0042-6989(98)00012-1

Figure/Ground Segregation from Temporal Delay is Best at High Spatial Frequencies.

Haruyuki Kojima*

Department of Psychology
Vanderbilt University
Nashville, TN 37240, USA

* to whom correspondence should be addressed

e-mail: kojima@ctrvax.vanderbilt.edu

RUNNING HEAD: Temporal Figure Segregation

KEYWORDS: figure-ground, temporal difference, spatial frequency, segregation,
binding

ABSTRACT

Two experiments investigated the role of spatial frequency in performance of a figure/ground segregation task based on temporal cues. Figure orientation was much easier to judge when figure and ground portions of the target were defined exclusively by random texture composed entirely of high spatial frequencies. When target components were defined by low spatial frequencies only, the task was nearly impossible except with long temporal delay between figure and ground. These results are inconsistent with the hypothesis that M-cell activity is primarily responsible for figure/ground segregation from temporal delay. Instead, these results point to a distinction between temporal integration and temporal differentiation. Additionally, the present results can be related to recent work on the binding of spatial features over time.

INTRODUCTION

One of visual perception's chief job is the segregation of objects from their backgrounds. To accomplish figure/ground segregation, visual nervous system must register the presence of edges and contours defining objects. We typically think of these contours being represented by discontinuities in luminance contrast, texture, color, motion or disparity, relative to the background (e.g. Julesz, 1984; Ramachandran & Anstis, 1986; Nakayama, Shimojo & Silverman, 1989; Nothdurft, 1991). However, figure contours can also be defined by temporal disparity alone, in absence of any other spatial cues. Under optimal conditions, a time difference as brief as 5 msec between onset of figure and the ground is sufficient to promote figural discrimination (Fahle, 1993). Thinking about the neural bases for this perceptual ability, it is natural to conclude that the underlying mechanism must be one with excellent temporal acuity, for otherwise temporal integration would "blur" figure and ground signals and, thereby, destroy their uniqueness. And only cells of magno-pathway would seem to possess the temporal resolution to support such acute psychophysical performance (for review, see e.g. Lennie, 1980; DeYoe & Van Essen, 1988; Merigan & Maunsell, 1993).

If M-cells are indeed the substrate for temporal figure discrimination, we would expect performance on such a task to be especially good at lower spatial frequencies where M-cells are highly responsive (e.g. Schiller & Logothetis, 1990; Schiller,

Logothetis & Charles, 1990; Merigan, Byrne & Maunsell, 1991). The present study tests this prediction by measuring the dependence of temporal segregation of figure and ground on spatial frequency. Contrary to expectations, I find that observers are much better at discriminating figures defined by temporal cues when those figures and their backgrounds are composed of high spatial frequencies only.

METHODS

Figure 1 about here

Stimuli were generated under control of Power Macintosh 7200/120 and presented on a 21 inch multi-sync monitor (NEC XE21), with 75Hz frame rate. The test stimulus consisted of a rectangular "target figure" appearing within a larger, square "ground" region. The ground subtended 4.25×4.25 deg and the rectangular target figure subtended 1.96×1.20 deg. The target figure could be oriented vertically or horizontally, always centered within the ground region. Both the target figure and the ground were textured with uniform random dots (Fig. 1 a, b; OR). Each dot size corresponding to one pixel subtended 1.14 arcmin and the dot density of the figure and the ground was 1130 dots/deg² (50 %). The luminance of the white part of the figure was 15.2 cd/m² and that of the black dots was 0.08 cd/m².

Three types of textured patterns were tested (Fig. 1); unfiltered random dot patterns (original: OR), patterns containing only low spatial frequencies (lowpass: LP) and patterns containing only high spatial frequencies (highpass: HP). Both the LP and

the HP patterns were made from the OR pattern with a spatial frequency filter whose cutoff frequency was 2.5 c/deg for LP, and with a filter which eliminates spatial frequencies below 8.4 c/deg for HP.

The space-average luminance of the figure and of the ground was 7.64 cd/m², irrespective of their spatial frequency control. When the ground was presented without its textured figure, the central rectangular region could appear either white (15.2 cd/m²) (Fig. 1 a) or gray (the same mean luminance of the figure itself, i.e. 7.64 cd/m²) (Fig. 1 d). When the figure was presented without its textured ground, the empty surround region could be also filled with either white (Fig. 1 b) or gray (Fig. 1 e). When a figure and the ground were presented with a extremely short temporal delay or without a delay, i.e. when they were superimposed (Fig. 1 c), observer can not see any central figure; a target figure can only be seen with a certain temporal delay. In the present experiments, the temporal delay between onset of figure and ground could be produced to an accuracy of 13.3 msec.

The experimental sessions were run in a darkened room. Subjects viewed the display with natural pupils from a distance of 105 cm and initiated trials at their own pace. During each 1 sec stimulus presentation, the central figure was oriented either horizontally or vertically - the subject's task was to judge which, guessing if necessary. In the first experiment, the delay between figure and ground was set constant to the minimum frame rate, 13.3 msec in 75 Hz frame rate, with the cycle rate of the figure/ground sequence varied by the range of 3, 5, 7.5, 15, or 25 Hz (Fig. 2 a). In the second experiment, the temporal delay defining figure and ground was varied parametrically within cycle rate (Fig. 2 b). Only the OR and LP conditions were tested in Experiment 2, for the results of Experiment 1 showed that HP targets were vary easy to discriminate under the most demanding conditions.

Figure 2 about here

For a given experiment, trials for all conditions were randomly intermixed with 40 trials per condition per subject. Four adults with normal or corrected to normal vision served as subjects for both experiments.

RESULTS

The results from Experiment 1 are shown in Fig. 3, which plots the percent correct performance as a function of the cycle rate of the stimulus figure; the color of the empty region is the parameter. 50 % correct corresponds to chance performance on this 2 AFC task. The pattern of results from all four observers was the same, so the graphs show averages with standard errors.

Figure 3 about here

When the empty figure and ground regions were filled white, the figure/ground segregation task was trivially easy for all three pattern conditions. This result merely confirm that luminance-defined objects readily segregate from a background even at very brief temporal delays. When the region was the same average luminance as the pattern, the ease of the task depended on the spatial frequency composition of the display. The performance was easy at all temporal frequencies with HP patterns; the task was difficult at low temporal frequencies, 3 and

5 Hz, for OR; and the task was very difficult to impossible over all temporal frequencies for LP patterns.

Figure 4 about here

Results from Experiment 2 are shown in Fig. 4. For the OR patterns, one additional frame of temporal delay between the figure and the ground served dramatically to improve performance. For the LP condition, however, performance improved only gradually with increasing temporal delays. For subjects to perform the figure/ground segregation task perfectly with LP stimuli, temporal delays in excess of 50 msec were required.

DISCUSSION

The present experiments demonstrate that figure/ground segregation based on temporal delay depends on the spatial frequency composition of the targets. Contrary to intuition, the presence of low spatial frequencies actually hinders performance on such a task. Thus observers found the task nearly impossible when only low spatial frequencies were present (LP)¹, somewhat difficult when low and high spatial frequencies were present (OR) and easy when only high spatial frequencies were present (HP). These results lead to several interesting conclusions.

The generally superior performance when targets contain only high spatial frequencies runs counter to expectations based on the spatio-temporal sensitivity of human vision. Work dating back to the mid '60s demonstrates that human vision

exhibits high temporal sensitivity for low spatial frequencies and low temporal sensitivity for high spatial frequencies (Robson, 1966; Tolhurst, 1973, 1975). This pattern of results provided the backbone for a large body of literature implicating sustained and transient mechanisms in human vision (e.g. Breitmeyer, 1984; Watson, 1986). In recent years, this sustained/transient dichotomy has given way to the more contemporary M-pathway/P-pathway distinction (Livingstone & Hubel, 1987; DeYoe & Van Essen, 1988; Schiller & Logothetis, 1990; Merigan & Maunsell, 1993), but the basic idea remains the same: temporal sensitivity is best at low spatial frequencies.

How can the present results be reconciled with this almost paradigmatic view of human spatio-temporal vision? Perhaps the answer turns on the distinction between temporal integration and temporal differentiation. When contours such as the bars of a grating flicker rapidly, contrast energy is summed within the limits of temporal integration. In human vision, the integration contrast is shorter at lower spatial frequencies, promoting greater flicker sensitivity (e.g. Kelly, 1966). But for figure segregation based on temporal delay, contrast energy within neighboring spatial regions must be differentiated. In human vision, the sharpness of spatial differentiation is greater at higher spatial frequencies with high contrast stimuli (Sagi & Hochstein, 1985). Consistent with this distinction between integration and differentiation, evidence shows that temporal discrimination of neighboring targets is better than temporal discrimination of the targets itself. Thus, observers can detect temporal differences in the onsets of adjustment lines with onset delays as brief as a few milliseconds (Westheimer and McKee, 1977). Yet a single line flickering at the equivalent rate would appear steady. Thus the temporal limits for figure/ground segregation as revealed by Experiment 1 are not set by the temporal resolution of the

visual system's M-pathway. Those limits are imposed by the mechanisms responsible for spatio-temporal differentiation.

Finally, the present results and conclusion may have some bearing on recent psychophysical work motivated by so-called binding problem. Several research groups have investigated whether temporal factors can enhance figure perception, reasoning that such an effect would be consistent with the "temporal oscillation" hypothesis advanced by physiologists (Gray, König, Engel & Singer, 1989; Engel, König, Kreiter & Singer, 1991; Singer & Gray, 1995). The results from psychophysical investigations, however, have been inconsistent. Kiper, Gegenfurtner and Movshon (1996) found that figure/ground for texture-defined objects was insensitive to temporal phase differences among those objects. Similarly, Fahle and Koch (1995) reported that temporal phase had no effect on the completion of extended boundaries over space. On the other hand, Leonards, Singer and Fahle (1996) found that temporal cues were effective in promotion of figural contours under conditions where spatial cues were ineffective. This finding squares with my results showing that figure/ground segregation was trivially easy under all conditions of temporal presentation when the two components of the display were defined by luminance contours. Only when spatially defined contours are weakened or eliminated do the effects of temporal cues become salient.

Acknowledgment

The author gratefully thanks Prof. Randolph Blake for his extensive comments and writing suggestions on the manuscript and anonymous reviewers for their comments. This research was supported by JSPS Fellowship for Research Abroad, 1996.

REFERENCES

- Breitmeyer, B. G. (1984). *Visual masking: an integrative approach*. Oxford, New York.
- DeYoe, E. A. & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neuroscience*, **11**, 219-226.
- Engel, A. E., König, P., Kreiter, A. K. & Singer, W. (1991). Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science*, **252**, 1177-1179.
- Fahle, M. (1993). Figure-ground discrimination from temporal information. *Proceedings of Royal Society, London*, **B, 254**, 199-203.
- Fahle, M. & Koch, C. (1995). Spatial displacement, but not temporal asynchrony, destroys figural binding. *Vision Research*, **35**, 491-494.

- Gray, C. M., König, P., Engel, A. K. & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334-337.
- Julesz, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neuroscience*, **7**, 41-45.
- Kiper, D. C., Gegenfurtner, K. R. & Movshon, A. J. (1996). Cortical oscillatory responses do not affect visual segmentation. *Vision Research*, **36**, 539-544.
- Kelly, D. H. (1966). Frequency doubling in visual responses. *Journal of the Optical Society of America*, **56**, 1628-1633.
- Lennie, P. (1980). Parallel visual pathways: A review. *Vision Research*, **20**, 561-594.
- Leonards, U., Singer, W. & Fahle, M. (1996). The influence of temporal phase differences on texture segmentation. *Vision Research*, **36**, 2689-2697.
- Livingstone, M. S. & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, **7**, 3416-3468.
- Merigan, W. H., Byrne, C. E. & Maunsell, H. R. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience*, **11**, 3422-3429.
- Merigan, W. H. & Maunsell, H. R. (1993). How parallel are the primate visual pathways? *Annual Review of Neuroscience*, **16**, 369-402.
- Nakayama, K., Shimojo, S. & Silverman, G. H. (1989). Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception*, **18**, 55-68.
- Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, **31**, 1073-1078.

- Ramachandran, V. S. & Anstis, S. (1986). Figure-ground segregation modulates apparent motion. *Vision Research*, **26**, 1969-1975.
- Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system, *Journal of the Optical Society of America*, **56**, 1141-1142.
- Sagi, D. & Hochstein, S. (1985). Lateral inhibition between spatially adjacent spatial-frequency channels? *Perception & Psychophysics*, **37**, 315-322.
- Schiller, P. H. & Logothetis, N. K. (1990). The colour-opponent and broad-band channels of the primate visual system. *Trends in Neuroscience*, **13**, 392-398.
- Schiller, P. H., Logothetis, N. K. & Charles, E. R. (1990). Functions of the colour-opponent and broad-band channels of the visual system. *Nature*, **343**, 68-70.
- Singer, W. & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, **18**, 555-586.
- Tolhurst, D. J. (1973). Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *Journal of Physiology*, **231**, 385-402.
- Tolhurst, D. J. (1975). Transient and sustained channels in human vision. *Vision Research*, **15**, 1151-1155.
- Watson, A. B. (1986). Temporal sensitivity. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds.), *Handbook of perception and human performance, vol. I Sensory processes and perception* (pp. 6-1 - 6-43). New York: John Wiley & Sons.
- Westheimer, G. & McKee, S. P. (1977). Integration regions for visual hyperacuity. *Vision Research*, **17**, 89-93.

FIGURE CAPTIONS

Fig. 1 Each diagrams illustrate the pattern conditions used in the experiments. (a) The "Ground" with a white empty region, (b) the target "Figure" with a white empty region, (c) the superimposed display of the "figure" and the "ground", (d) the "Ground" with an empty region filled with gray equal to the mean luminance of the pattern, (e) the target "Figure" with the gray. Each rows shows High Pass patterns (HP), Original Random dot patterns (OR) and Low Pass patterns (LP), respectively.

Fig. 2 Schematic temporal sequence of stimulus presentation for (a) the experiment 1 and for (b) the experiment 2. As the figure/ground was defined only by the temporal delay between them, a figure and the ground were intervened in each stimulus cycles, otherwise presented “Superimposed” display as seen in Fig. 1. The temporal rate of monitor frames (F_i) was constant 75 Hz (the interval of each frames was 13.3 msec). The presentation cycles were repeated for 1 sec. In Experiment 1, cycle rate of figure/ground presentation, the number of the delayed frames within one second was varied. In Experiment 2, the number of frames inserted between figure display and the ground was varied, while the cycle rate of figure/ground was kept constant.

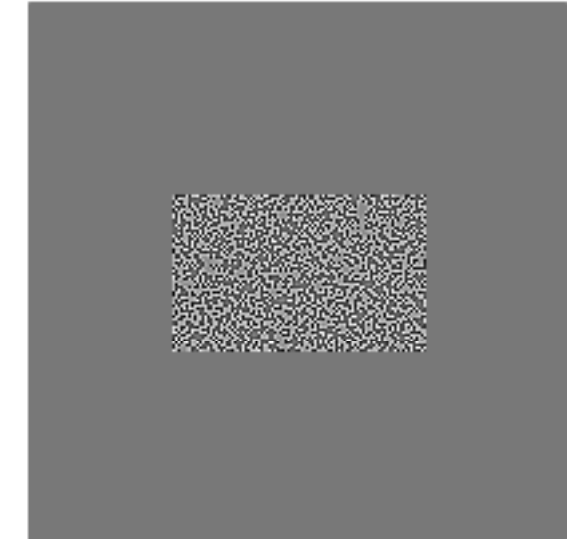
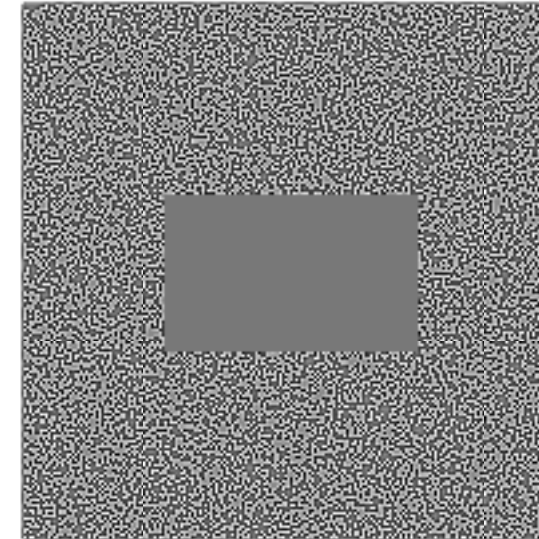
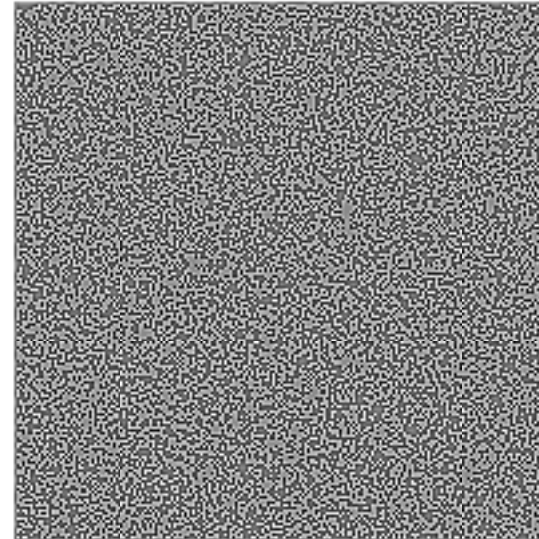
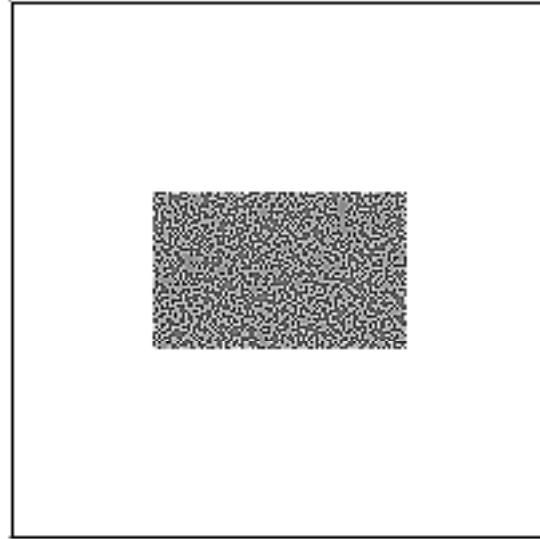
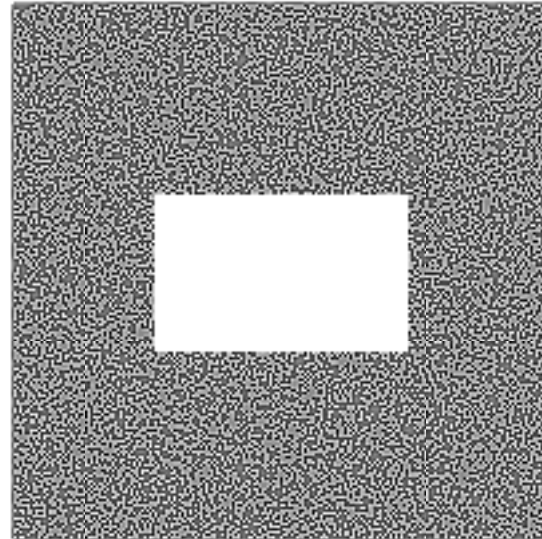
Fig. 3 Averages of performance, percent correct, in the experiment 1 were shown as a function of the cycle rate of stimulus (Hz) with standard errors of subjects. In some conditions, standard errors were too small to be seen. Each panels represents different pattern conditions; HP, OR and LP, respectively. For each panels, open squares show the performance for patterns filled with white, while solid circles represent those with mean luminance gray.

Fig. 4 Average performance (%) in Experiment 2 was plotted as a function of the temporal delay between figure and the ground (msec). Solid triangle, solid square, and solid diamond represent 3, 5, and 7.5 (Hz), respectively, whereas, in comparison, the performances of the same temporal conditions (Hz) in Experiment 1 were shown as the corresponding shapes with open symbols.

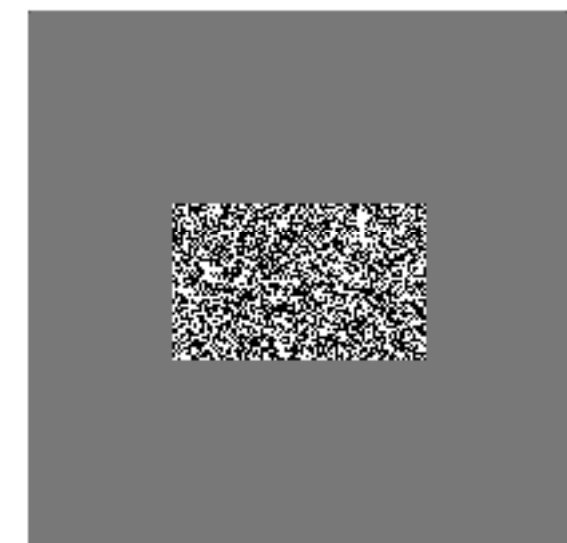
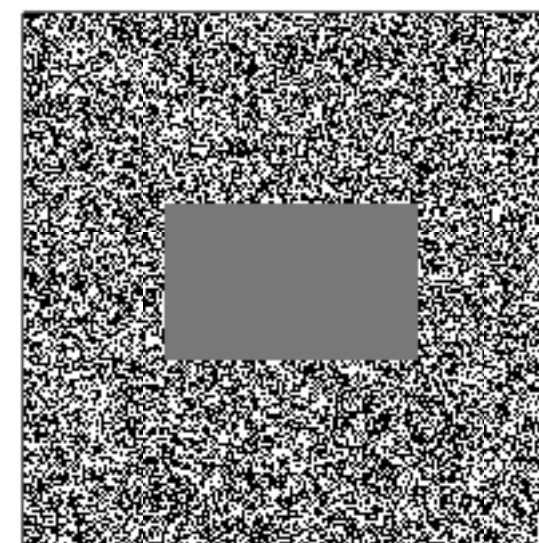
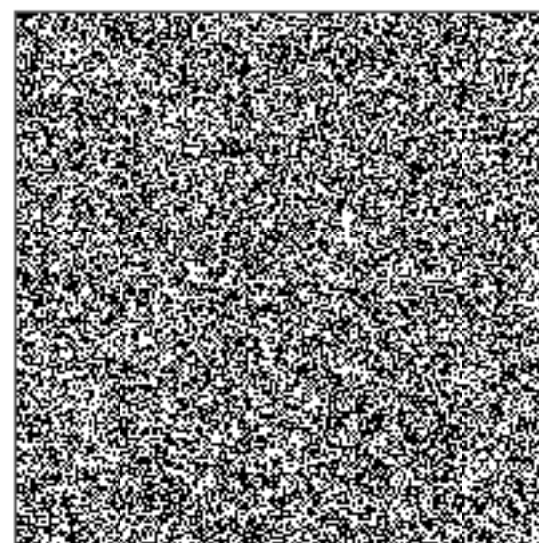
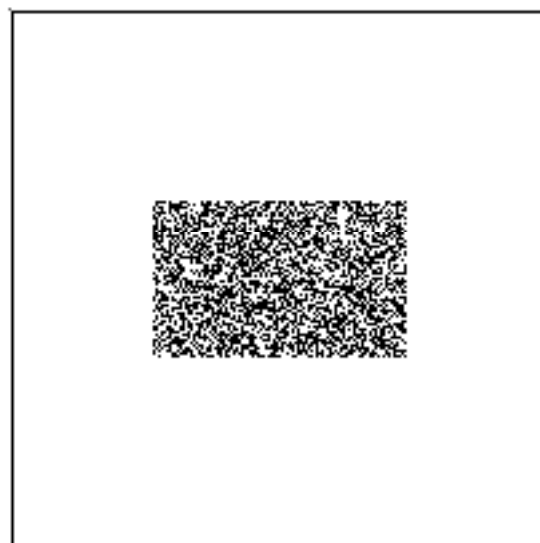
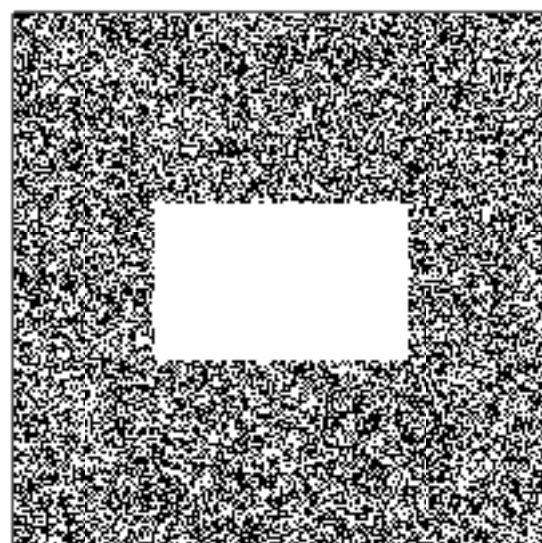
FOOTNOTES

1. The edge between LP figure and LP background was, in fact, sharp. Blurring this edge (thus eliminating the associated high spatial frequencies) would only make the task more difficult.

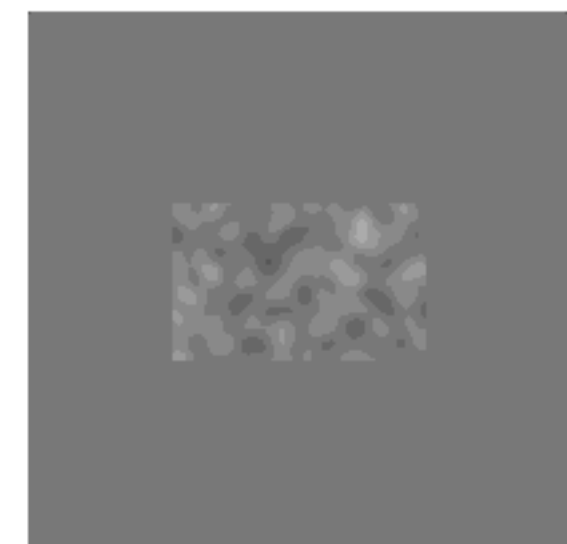
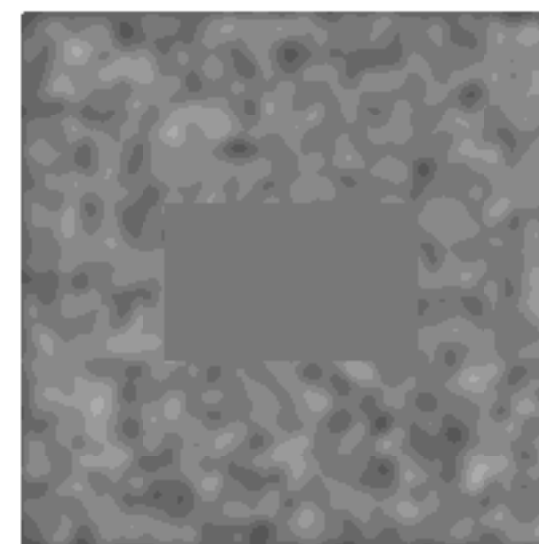
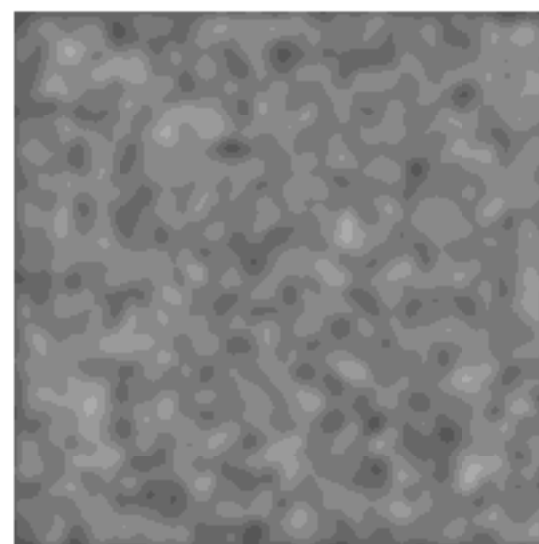
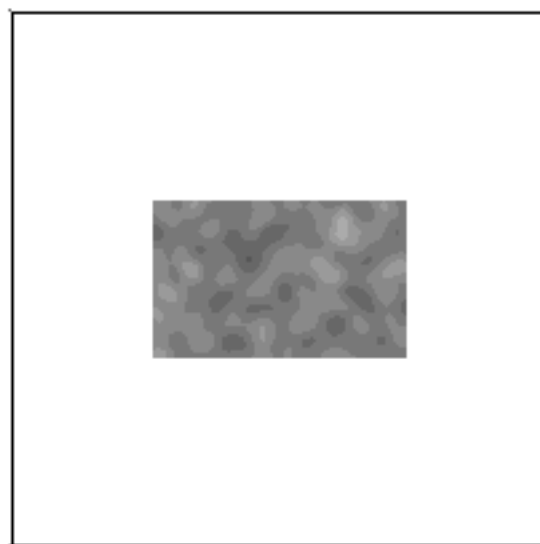
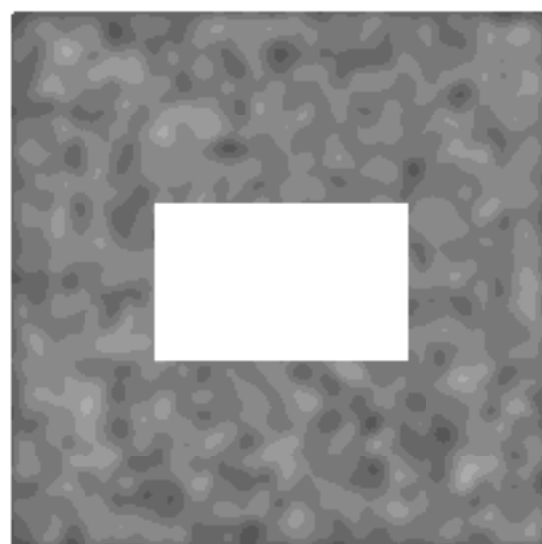
HP



OR



LP



(a) "Ground" filled
with White

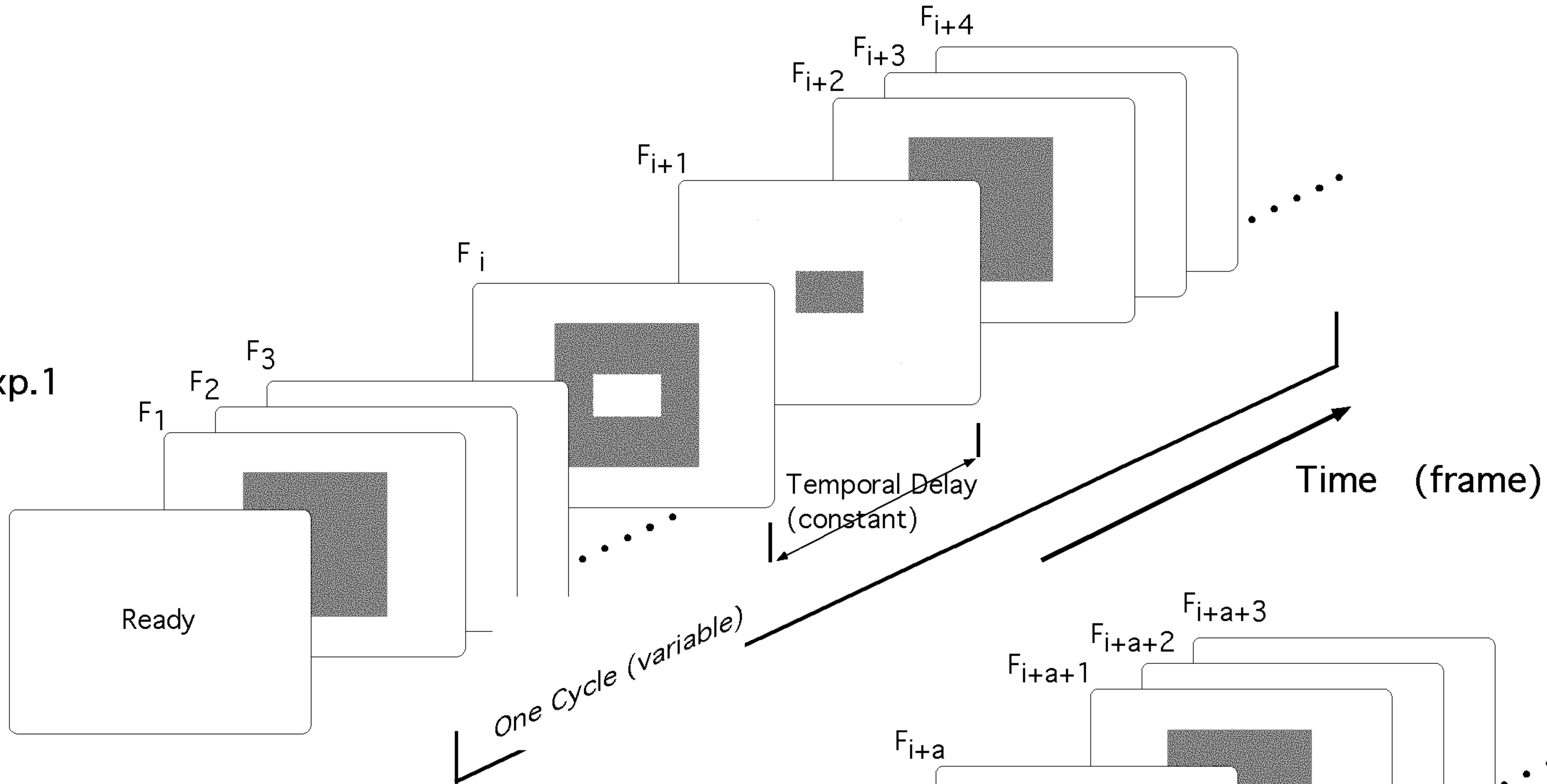
(b) "Figure" filled
with White

(c) "Superimposed"

(d) "Ground" filled
with Gray

(e) "Figure" filled
with Gray

(a) Exp.1



(b) Exp.2

